

EDdb: A web resource for eating disorder and its application to identify an extended adipocytokine signaling pathway related to eating disorder

ZHAO Min[†], LI XiaoMo[†] & QU Hong^{*}

Center for Bioinformatics, State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing 100871, China

Received January 2, 2013; accepted May 23, 2013

Eating disorder is a group of physiological and psychological disorders affecting approximately 1% of the female population worldwide. Although the genetic epidemiology of eating disorder is becoming increasingly clear with accumulated studies, the underlying molecular mechanisms are still unclear. Recently, integration of various high-throughput data expanded the range of candidate genes and started to generate hypotheses for understanding potential pathogenesis in complex diseases. This article presents EDdb (Eating Disorder database), the first evidence-based gene resource for eating disorder. Fifty-nine experimentally validated genes from the literature in relation to eating disorder were collected as the core dataset. Another four datasets with 2824 candidate genes across 601 genome regions were expanded based on the core dataset using different criteria (e.g., protein-protein interactions, shared cytobands, and related complex diseases). Based on human protein-protein interaction data, we reconstructed a potential molecular sub-network related to eating disorder. Furthermore, with an integrative pathway enrichment analysis of genes in EDdb, we identified an extended adipocytokine signaling pathway in eating disorder. Three genes in EDdb (*ADIPO* (adiponectin), *TNF* (tumor necrosis factor) and *NR3C1* (nuclear receptor subfamily 3, group C, member 1)) link the KEGG (Kyoto Encyclopedia of Genes and Genomes) “adipocytokine signaling pathway” with the BioCarta “visceral fat deposits and the metabolic syndrome” pathway to form a joint pathway. In total, the joint pathway contains 43 genes, among which 39 genes are related to eating disorder. As the first comprehensive gene resource for eating disorder, EDdb (<http://eddb.cbi.pku.edu.cn>) enables the exploration of gene-disease relationships and cross-talk mechanisms between related disorders. Through pathway statistical studies, we revealed that abnormal body weight caused by eating disorder and obesity may both be related to dysregulation of the novel joint pathway of adipocytokine signaling. In addition, this joint pathway may be the common pathway for body weight regulation in complex human diseases related to unhealthy lifestyle.

eating disorder, database, adipocytokine signaling pathway, pathway analysis

Citation: Zhao M, Li X M, Qu H. EDdb: A web resource for eating disorder and its application to identify an extended adipocytokine signaling pathway related to eating disorder. *Sci China Life Sci*, 2013, 56: 1086–1096, doi: 10.1007/s11427-013-4573-2

Eating disorder (MeSH D001068) is a group of physiological and psychological disorders defined by abnormal diet behavior. It mainly consists of several diagnoses: anorexia nervosa (AN), bulimia nervosa (BN), binge eating disorder

(BED) and eating disorder not otherwise specified (EDNOS) [1,2]. Besides their disturbances in appetite and food intake, the irrational attitude towards weight and body shape may cause suicidal thinking. Interestingly, women are affected 10 times more frequently than men [3]. It has been estimated that nine in 1000 women in the USA are affected by anorexia nervosa [4]. However, current genetic predisposition

[†]Contributed equally to this work

^{*}Corresponding author (email: quh@mail.cbi.pku.edu.cn)

information is scattered in the published literature. Only 12 genes with positive evidence are stored in GAD (Genetic Association Database) [5]. A more comprehensive gene resource needs to be developed to acquire a more complete picture of eating disorder.

Common unhealthy lifestyles, such as smoking, can increase the risk of several metabolic disorders including eating disorder, hypertension, obesity, type 2 diabetes mellitus and ischemic heart disease [6–9]. In particular, the symptoms of eating disorder often overlap with other weight-related disorders such as obesity because of common unhealthy diets [10–16]. Although the cause of obesity and eating disorder is becoming increasingly clearer with large-scale and individual genetic studies, the common molecular pathway underlying abnormal body weight between these two weight-related disorders has not been analysed yet.

Although genome-wide association studies have revealed a number of loci in complex diseases such as obesity [17,18], two high-throughput studies on eating disorder were unsuccessful in revealing more loci. One study was designed to scan genomic regions of 182 genes from the literature, but resulted in no significant SNPs (single-nucleotide polymorphisms) for anorexia nervosa [19]. The other genome-wide association study on anorexia nervosa sampled 4766 individuals, but confirmed only one common SNP within *OPRD1* (Opioid Receptor, Delta 1) [3].

Systematic and statistical analysis of genes and their pathways may provide a more complete pathway related to the molecular mechanism underlying eating disorder. The majority of identified genes related to eating disorder can be classified into two groups, which are relevant to the neurotransmitter system and body weight regulation. For neurotransmitters, several systems have been proposed, such as the serotonergic system, noradrenergic system and dopaminergic system, while genes for weight regulation are mainly associated with the leptinergic-melanocortinergic system [20]. The association of the molecular mechanisms of these systems with other relevant diseases remains to be explored.

In this article, we developed an eating disorder gene database, EDdb (Eating Disorder database), for cataloguing genes linked to this disorder. Our data representing genetic knowledge encompasses a collection of all types of evidence including mutation analysis, gene expression, gene association, family-based linkage studies, genome-wide association studies and other functional studies. The focus of this resource is to provide an up-to-date and annotated list of eating disorder candidate genes to serve as reference datasets for understanding its etiology. In addition, we set up a strategy for unveiling novel eating disorder susceptibility genes by protein-protein interactions, shared cytobands or related diseases such as obesity. Based on integrated protein-protein interaction data, we have reconstructed a

potential network closely related to eating disorder. Furthermore, our systematic and statistical pathway studies revealed that the abnormal body weight caused by eating disorder may be caused by dysregulation of important signaling events occurring in a joint pathway combining the KEGG (Kyoto Encyclopedia of Genes and Genomes) “adipocytokine signaling pathway” and the BioCarta “visceral fat deposits and the metabolic syndrome” pathway.

1 Materials and methods

1.1 Data collection and automation of gene annotation

The genes related to eating disorder in the core dataset were extracted from 567 abstracts from the year 2010, retrieved by PubMed query ‘((anorexia nervosa) OR (bulimia nervosa) OR (eating disorder)) AND gene’. The data spanned multiple technology platforms including single-gene mutation analysis, family-based linkage studies, population association studies, gene expression analysis, genome-wide association studies and other functional analyses. Seven raw datasets were prepared for the gene list expansion. The first raw dataset included four controversial genes (*DRD3*, *TNF*, *UCP2* and *UCP3*) with equally positive and negative evidence from individual articles. Another literature dataset was collected from the supplement of Pinheiro et al. [19]. In addition, three datasets of related diseases (obesity, bipolar disorder and schizophrenia) were collected from Gene Association Database [5]. For accuracy, we excluded negative records of the three disorders in Gene Association Database. The protein-protein interaction partners of genes in the core dataset were integrated from the BioGRID (September 25, 2011 Version 3.1.81) [21], HPRD (April 13, 2010 Version 9) [22], and BIND (downloaded on December 15, 2009) [23] databases. Finally, the last raw dataset contained genes that shared cytobands with genes from the core dataset [24].

The core dataset will be routinely updated based on newly published literature. In addition to continuous data curation and integration of literature and other disease related databases, we will focus on constructing regulatory networks in humans with emphasis on their regulatory transcription factors and phosphorylation. An automatic annotation pipeline was implemented to integrate functional information about the genes from the Entrez gene database (January 13, 2011) [24], Gene Ontology annotation, interacting proteins from HPRD/BIND/BioGRID, and OMIM annotation using Perl script and Swiss knife module [22,25–27]. Signaling event annotation was added to each gene according to the KEGG LIGAND and BioCarta databases [28,29]. With the automatic pipeline for data annotation, our database can be easily updated when new versions of external databases such as BioCarta and KEGG LIGAND are available.

1.2 Sub-network extraction

To obtain a network of eating disorder genes, we firstly constructed a non-redundant human interactome based on the protein-protein interaction data from the HPRD, IntAct, BioGRID and MINT databases. In addition, we excluded self-interactions, interactions predicted by computational methods, and interactions with other species for five networks. The final interactome contained 12363 nodes and 70019 protein-protein interaction links. The degree of all nodes in this human PPI network followed the power law distribution, $P(k) \sim k^{-b}$, where $P(k)$ represents the probability that a node interacts with k nodes and b is an exponent with an estimated value of 1.683. This indicated that most genes in this human PPI network were sparsely connected, while a small part of them had relatively high connectivity. Using the Steiner minimal tree algorithm [30], we extracted the sub-network as an eating disorder-specific network from the whole human interactome. In our implemented algorithm, the sub-network was first constructed with all input genes in the core dataset connected. Finally, a minimum net with input genes connected by shortest path was produced.

1.3 Web interface construction

All data and information in EDdb are stored in a MySQL relational database on a Linux server [31–33]. The database provides a web interface, which allows researchers to browse candidate genes from different features. Web-based queries to the database were implemented in Perl scripts running in an Apache environment. EDdb allows users to do text query, or to run BLAST searches against the sequences in EDdb. For an advanced study, this database provides all nucleotide and protein sequences in a FASTA format for all the entries. Furthermore, cross-references to external data-

bases such as KEGG and iHOP (information hyperlinked over proteins) are included for each record in EDdb. In addition, for each gene in the core dataset, links to the original articles related to eating disorder in the NCBI PubMed database are given (Figure 1).

1.4 Identification of representative pathways using KOBAS

Throughout the paper, the representative pathways from the KEGG and BioCarta databases for each gene set were identified by KOBAS (KEGG Orthology Based Annotation System) [34]. In these pathway analyses, all human protein-coding genes in KOBAS were set as the background to calculate statistical significance. In addition, the false discovery rate (FDR)-corrected P -values for enriched pathways were adopted based on the hypergeometric test. Finally, the enriched human pathways with corrected P -values of less than 0.01 were identified as over-representative pathways for each gene set.

1.5 Empirical resampling for evaluating the relationships between eating disorder and other diseases

To determine whether eating disorder is related to obesity, schizophrenia and bipolar disorder more than expected, we conducted a total of 1000 resamplings. Here, we took the obesity gene for example. During each process, 59 human genes were randomly selected from 45522 human genes in the Entrez gene database (January 13, 2011). We repeated this random process 1000 times.

Next, we counted the number of randomly selected node sets (N) whose number of overlapping genes was more than the actual number of overlapping genes [33]. Finally, we

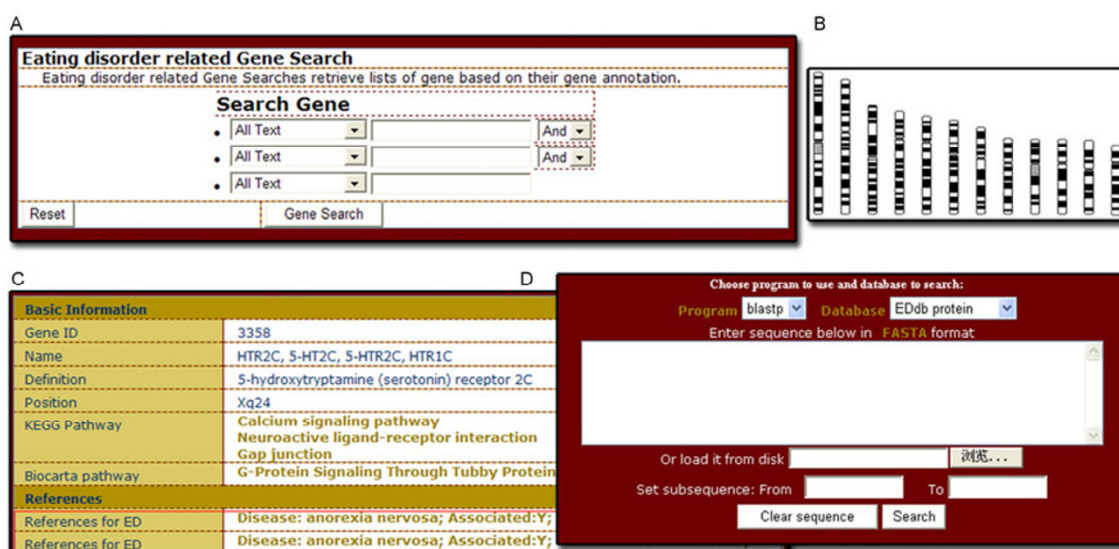


Figure 1 Web interface of EDdb. A, Query interface for text search. B, Data browser by chromosomes. C, The related literature contents in each confirmed gene page. D, Blast interface for sequence search against all genes in EDdb.

calculated the empirical *P*-value using $N/1000$ to represent the significance of how the 59 eating disorder genes are related to obesity in gene content. Next, we applied a similar approach to schizophrenia and bipolar disorder. Based on this empirical resampling approach, the empirical *P*-values for the relationship between the eating disorder gene set and obesity, schizophrenia and bipolar disorder are all less than 0.01.

2 Results

As the first literature-based data resource for eating disorder, our database provides clearer clues for classifying this disease and reveals an underlying biological pathway model from a genetic view.

2.1 Collection of the core dataset, consisting of experimentally verified candidate genes for eating disorder

As shown in Figure 2, the comprehensive collection of experimentally verified genes in EDdb was accomplished by curating published literature in the following four steps: (i) An extensive literature search was conducted through PubMed with the keywords “eating disorder”, “anorexia nervosa”, “bulimia nervosa” or “binge eating disorder” (see Materials and methods); (ii) the retrieved abstracts were highlighted with query keywords and grouped by the Related Articles function in Entrez; (iii) the abstracts were manually read to list the experimentally verified candidate genes for eating disorder and other related information such as experimental methods and populations; (iv) the candidate gene names were manually mapped to Entrez Gene ID and automatic annotation was performed (see Materials and methods). We also recorded negative results for each gene, and we assumed that genes with more positive reports were candidate genes. Only four genes (*DRD3* (Dopamine Receptor D3), *TNF* (tumor necrosis factor), *UCP2* (uncoupling protein 2) and *UCP3* (uncoupling protein 3)) with equal negative and positive reports were excluded from the core dataset. Finally, 59 confirmed genes related to eating disorder were stored in this core dataset, including three overlapping subsets: 52 genes related to anorexia nervosa, 19 genes related to bulimia nervosa, and 25 genes related to eating disorder. Among these 59 genes, 33, 31 and 20 were related to obesity, schizophrenia and bipolar disorder, respectively. Further empirical re-sampling demonstrated that eating disorder was closely related to obesity, schizophrenia and bipolar disorder in genetics, which were all significantly related to the 59 genes related to eating disorder (see section 4.5).

So far gene-set enrichment analysis is probably the most practical and successful data mining approach for addressing the challenge of analyzing complex gene lists [35]. Here, we employed KEGG and BioCarta pathway enrichment

analysis to summarize the possible signaling mechanisms of eating disorder. As listed in Table 1, the top 10 enriched pathways were mostly related to endocrine and nervous system signaling pathways, amino acid metabolism as well as fat deposit and metabolism. The statistical enrichments of amino acid metabolism could be explained by the *MAOA* (monoamine oxidase A) and *MAOB* (monoamine oxidase B) genes, which are involved in the metabolism of several amino acids tyrosine, tryptophan, phenylalanine and histidine. Because of their vital roles in the inactivation of neurotransmitters, both genes were thought to be related to neural disorders such as Parkinson’s disease [36,37].

2.2 Expanding candidate genes by protein-protein interactions and related diseases

The above mentioned systematic collection had limited success in discovering confirmed causal genes of eating disorder. Earlier studies showed that protein-protein interaction network, gene co-expression network and pathway enrichment analysis may provide a practical approach to expand the gene list and unveil the underlying molecular mechanisms [38]. We therefore set up a pathway discovery strategy to expand signaling mechanisms behind eating disorder combining protein-protein interactions, gene cytobands and related diseases. Seven raw datasets were prepared for the gene list expansion, including two datasets from the literature (see Materials and methods), three related disease (obesity, bipolar disorder and schizophrenia) datasets from Gene Association Database [5], one interaction dataset (genes interacting with genes from the core dataset) and one cytoband dataset (genes sharing cytobands with genes from the core dataset). First, we added 107 genes (hereafter referred to as EXP1 dataset) from the two literature datasets as they were also present in the other five related raw datasets. Next, for the second expansion (hereafter referred to as EXP2 dataset) 148 genes were added from the three disease datasets, as they also had support from the interaction and cytoband datasets. With more than two direct functional confirmations, the 255 genes from the EXP1 and EXP2 datasets were convincingly important in the pathogenesis of eating disorder, making them very good candidates for further functional investigations. The third expansion was 793 genes (hereafter referred to as EXP3 dataset) with evidence from only three disease datasets. The last one was 1775 genes (hereafter referred to as EXP4 dataset) with support only from the protein interaction dataset.

Once the extension of candidate genes was achieved, a subsequent interpretation task was required to evaluate the biological roles of the expanded genes. Because most of the genes play a wide variety of roles in different cellular functions, great caution was taken to avoid introducing non-specific pathways to data interpretation. Our pathway context analysis was conducted using a hypergeometric statistical comparison. Each expansion was involved in several

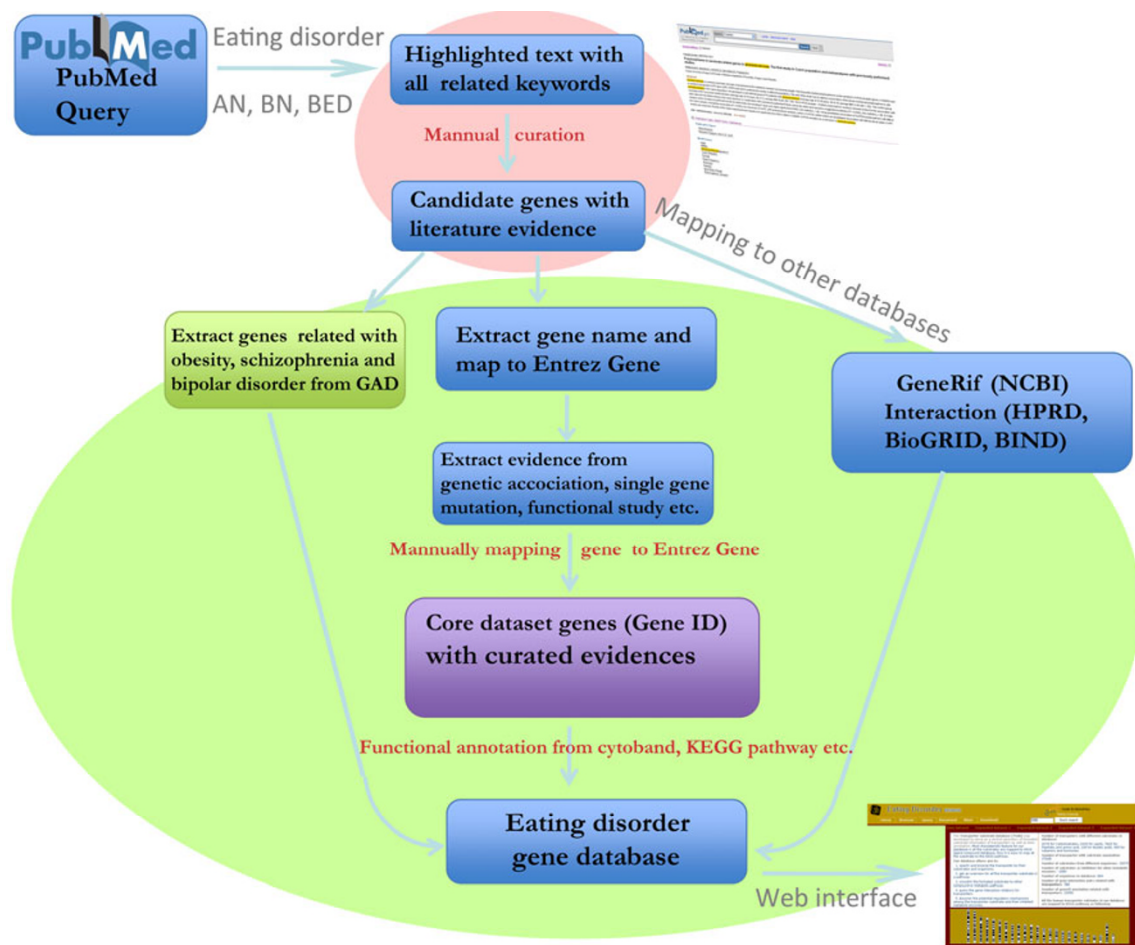


Figure 2 Pipeline for collection, expansion and annotation of genes related to eating disorder.

Table 1 Top 10 significantly enriched signaling pathways of 59 genes in the core dataset from the KEGG and BioCarta databases^{a)}

Pathway	Source	<i>P</i> -value ^{a)}
Neuroactive ligand-receptor interaction	KEGG	3.98×10^{-11}
Adipocytokine signaling pathway	KEGG	6.63×10^{-6}
Reversal of insulin resistance by leptin	BioCarta	5.52×10^{-4}
Metabolism of anandamide, an endogenous cannabinoid	BioCarta	7.69×10^{-4}
Visceral fat deposits and the metabolic syndrome	BioCarta	1.02×10^{-3}
Tyrosine metabolism	KEGG	2.22×10^{-3}
Tryptophan metabolism	KEGG	2.39×10^{-3}
Phenylalanine metabolism	KEGG	6.29×10^{-3}
Neurotrophin signaling pathway	KEGG	1.04×10^{-2}
Histidine metabolism	KEGG	1.59×10^{-2}

a) *, Corrected by Benjamin-Hochberg.

signaling pathways, including the calcium signaling pathway, long-term depression, Gap Junction and PPAR (peroxisome proliferator-activated receptor) signaling pathway (Table S1 in Supporting Information). Some of these pathways, e.g., Gap Junction and long-term depression, should perhaps be considered as generic rather than specific pathway, indicating the presence of general processes in neural disorders. Interestingly, only neuroactive ligand-receptor

interaction and the adipocytokine signaling pathway were always present in the top 10 pathways in the different datasets. They are intuitively linked with eating disorder, because of their close relationships with neural activity. However, in the KEGG pathway, the neuroactive ligand-receptor interaction was only a list of interaction pairs between neuroactive signaling molecules and their receptors instead of a fully connected map. Thus, it was very difficult

to achieve a connected molecular network. On the contrary, the adipocytokine signaling pathway is a fully connected graph. More interestingly, adipocytokines such as leptin and adiponectin are important signaling molecules involved in neuronal circuits for food intake and are related to obesity, insulin resistance and cancer [39–44].

2.3 Identifying representative genes in the adipocytokine signaling pathway related to eating disorder

Although leptin and adiponectin in the adipocytokine signaling pathway are related to anorexia nervosa and bulimia nervosa [45,46], other downstream signaling circuits have not been reported to be related to eating disorder. Based on our gene collection, most of the genes in the adipocytokine signaling pathway are related to eating disorder (Figure 3).

Six genes in the core dataset, including *LEP* (leptin), *LEPR* (leptin receptor), *ADIPO* (adiponectin), *POMC* (proopiomelanocortin), *AGRP* (agouti related protein homolog) and *NPY* (neuropeptide Y), were annotated to the adipocytokine signaling pathway. *LEP*, *LEPR* and *ADIPO* are the most important upstream signals for food intake [42,47]. The remaining three genes, *POMC*, *AGRP* and *NPY*, also play important roles in downstream signals in weight regulation or food intake [48–51]. Because of equally nega-

tive and positive reports, *TNF* (tumor necrosis factor) was not included in the core dataset, but only in the EXP1 dataset. *TNF* interferes with the early steps of insulin signaling [52,53]. Moreover, three genes including *SOCS3* (suppressor of cytokine signaling 3), *PPARGC1A* (peroxisome proliferator-activated receptor gamma, coactivator 1 alpha) and *ACC2* (acetyl-CoA carboxylase beta) were included in the EXP1 dataset. Both *PPARGC1A* and *ACC2* are involved in the regulation of fatty acid oxidation [54,55]. Though *SOCS3* cannot directly influence fatty acid oxidation, it inhibits leptin activation of *AMPK* (5' AMP-activated protein kinase) [56]. In summary, 10 genes from the core dataset and the EXP1 dataset can be clustered into two categories: (i) four upstream signals including *TNF*, *LEP*, *LEPR* and *ADIPO*, which are related to obesity; (ii) six downstream effector molecules of metabolic regulation, among which five were also related to obesity. As genes from both the core dataset and the EXP1 dataset are shown to be involved in eating disorder, our next task was to fill in the gap between highly convincing upstream signals and downstream effector genes with data from protein-protein interactions and related disease studies.

As shown in Figure 3, five genes from EXP2 are colored in the map. Three of them are receptors for *ADIPO* and *TNF*. The remaining two genes, *STAT3* (signal transducer and

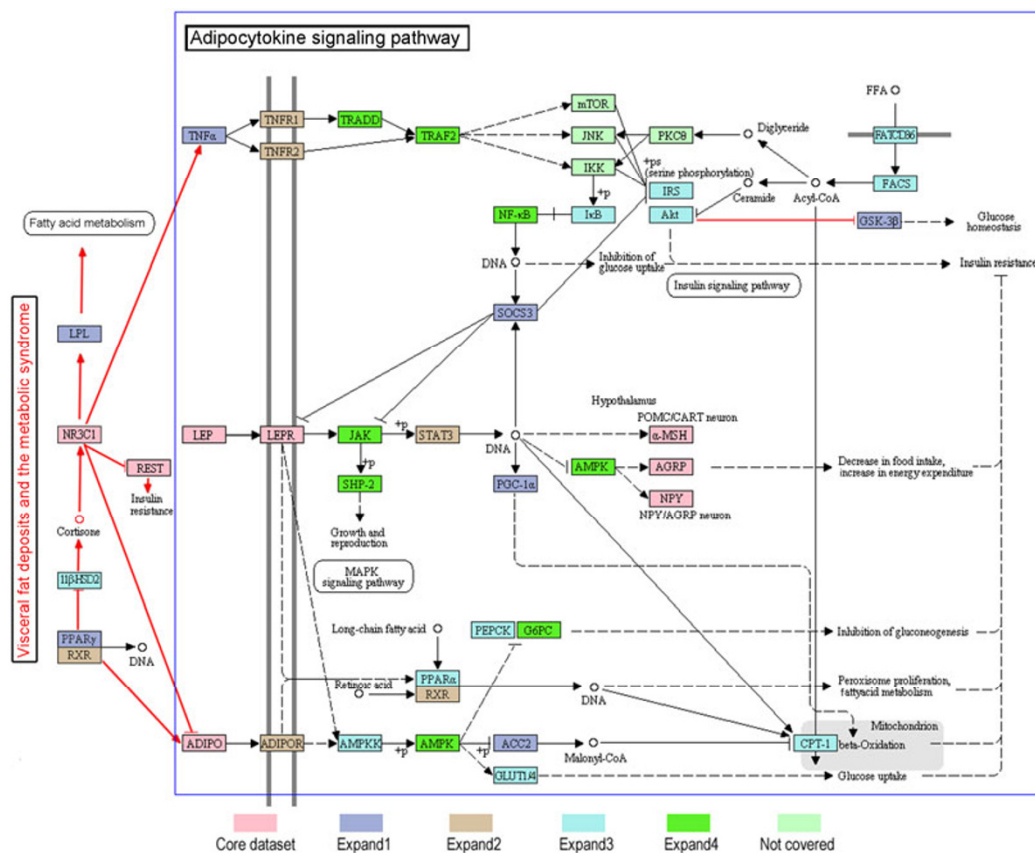


Figure 3 Key signaling map for body weight regulation in eating disorder combined from “visceral fat deposits and the metabolic syndrome” and “adipocytokine signalling pathway”. The genes from different datasets are color-coded. The color schemas for the different datasets are shown at the bottom of the map.

activator of transcription 3) and *RXR* (retinoid X receptor, alpha), are intermediate regulators in the adipocytokine signaling pathway. Furthermore, 10 genes from EXP3 and eight genes from EXP4 were introduced into the adipocytokine signaling pathway. Finally, all the gaps between upstream signals and downstream molecules were filled in the map with only four genes (*MTOR* (mechanistic target of rapamycin), *JNK* (c-Jun N-terminal kinase), *IKK* (I κ B kinase) and *PRKCQ* (Protein Kinase C, Theta)), which were not included in EDdb. Combining the four expansion datasets, we identified 33 genes in the adipocytokine signaling pathway, among which 20, 13 and 4 were associated with obesity, schizophrenia and bipolar disorder, respectively. In particular, the high proportion of obesity related genes in the adipocytokine signaling pathway may provide clues for explaining weight regulation in both diseases.

2.4 Key joint signaling map for eating disorder combined from the “adipocytokine signaling pathway” and “visceral fat deposits and the metabolic syndrome”

Different pathway databases have various pathway concepts [31,57]. The real cellular communications are dynamic and have no artificial boundaries. Connecting and combining common pathways from different databases through their cross-talking molecules is a direct approach to extend pathway paradigms. Our in-depth analysis of the top 10 statistically representative pathways in Table 1 revealed that the “reversal of insulin resistance by leptin” pathway from BioCarta is a complete subset of the “adipocytokine signaling pathway” from KEGG. Meanwhile, the “visceral fat deposits and the metabolic syndrome” pathway from BioCarta and the “adipocytokine signaling pathway” from KEGG can be combined together to form a fully connected map. “Visceral fat deposits and the metabolic syndrome” pathway contains eight molecules, among which three downstream or mid-stream genes (*ADIPO*, *TNF* and *RXR*) are the same upstream or mid-stream molecules found in the “adipocytokine signaling pathway”. The remaining five Upstream molecules include *PPARG* (peroxisome proliferator-activated receptor gamma), *HSD11B2* (hydroxysteroid 11-beta dehydrogenase 1), *NR3C1* (nuclear receptor sub-family 3, group C, member 1), *LPL* (lipoprotein lipase) and *REST* (resistin), among which *REST* and *NR3C1* have been confirmed to be related to eating disorder [58–60]. In addition, *PPARG* and *LPL* are both from the KEGG “PPAR signaling pathway”. Furthermore, at least five genes in Figure 3, including *LPL*, *PEPCK* (phosphoenolpyruvate carboxykinase), *CPT1A* (carnitine palmitoyltransferase 1A), *FACS* (acsl1 acyl-CoA synthetase long-chain family member 1) and *FAT/CD36* (fatty acid translocase/cluster of differentiation 36), are regulatory targets of *PPARG/RXR*, *PPARG/RXR* and *PPARA/RXR*. Thus, these genes form a highly linked regulatory module, bridging the “PPAR sig-

naling pathway” and the “adipocytokine signaling pathway”, hence highlighting the potential roles of these genes in the pathogenesis of eating disorder. In summary, 43 genes, mainly from the “visceral fat deposits and the metabolic syndrome” pathway and the “adipocytokine signaling pathway” are interlinked through regulatory relationships between *ADIPO*, *TNF* and *NR3C1*. Other genes in our database, such as *GSK3B* (glycogen synthase kinase 3 beta), from other relevant signaling pathways could also be connected to our summarized key joint signaling map (Figure 3). More importantly, 39 among the 43 genes are stored in our EDdb, and 25 of them are associated with obesity.

We observed a few more regulatory modules in the summarized key signaling map (Figure 3), for instance, the upstream regulatory module consisting of *ADIPO*, *PPARG/RXR*, *HSD11B2* and *NR3C1*, and another regulatory module, which is centered around *LEPR*, *SOC3* and *JAK* (Janus kinase). Cooperation of these coupled loops may transmit upstream signals to downstream cellular systems in fast or slow manners. Overall, our integration of several pathways from the BioCarta and KEGG databases revealed that eating disorder is a complex process caused by dysregulation of multiple highly interconnected pathways.

2.5 Reconstruction of a core pathway for eating disorder genes using sub-network extraction from protein-protein interaction data

The integrative pathway analysis above gave us a clue on how to reconstruct a core pathway for eating disorder. Here to reconstruct a potential biological process based on known protein-protein interactions in our EDdb, we firstly adopted a sub-network extraction approach. To get a comprehensive picture of eating disorder, we pieced together the protein-protein interaction data from the HPRD (Human Protein Reference Database), IntAct, BioGRID (Biological General Repository for Interaction Datasets) and MINT (Molecular INTERaction) databases. Based on 59 genes in our core dataset as seed, we used the Klein-Ravi algorithm to extract a sub-network from the integrated human protein-protein interaction (PPI) network (see Materials and methods). The final sub-network contains 78 genes in total (Figure 4). Fifty-three genes were in our eating disorder core dataset and the remaining 25 genes were linked to the 53 genes, forming a fully connected network. Functional enrichment analysis showed that the 78 genes in our reconstructed sub-network were enriched in six pathways (Table 2). The pathways “neuroactive ligand-receptor interaction”, “adipocytokine signaling pathway” and “visceral fat deposits and the metabolic syndrome” were the same as the enriched pathways in Table 1. The related pathways of three neurotransmitters including “serotonergic synapse”, “cocaine addiction”, and “dopaminergic synapse” were also detected. These results not only confirm the potential cen-

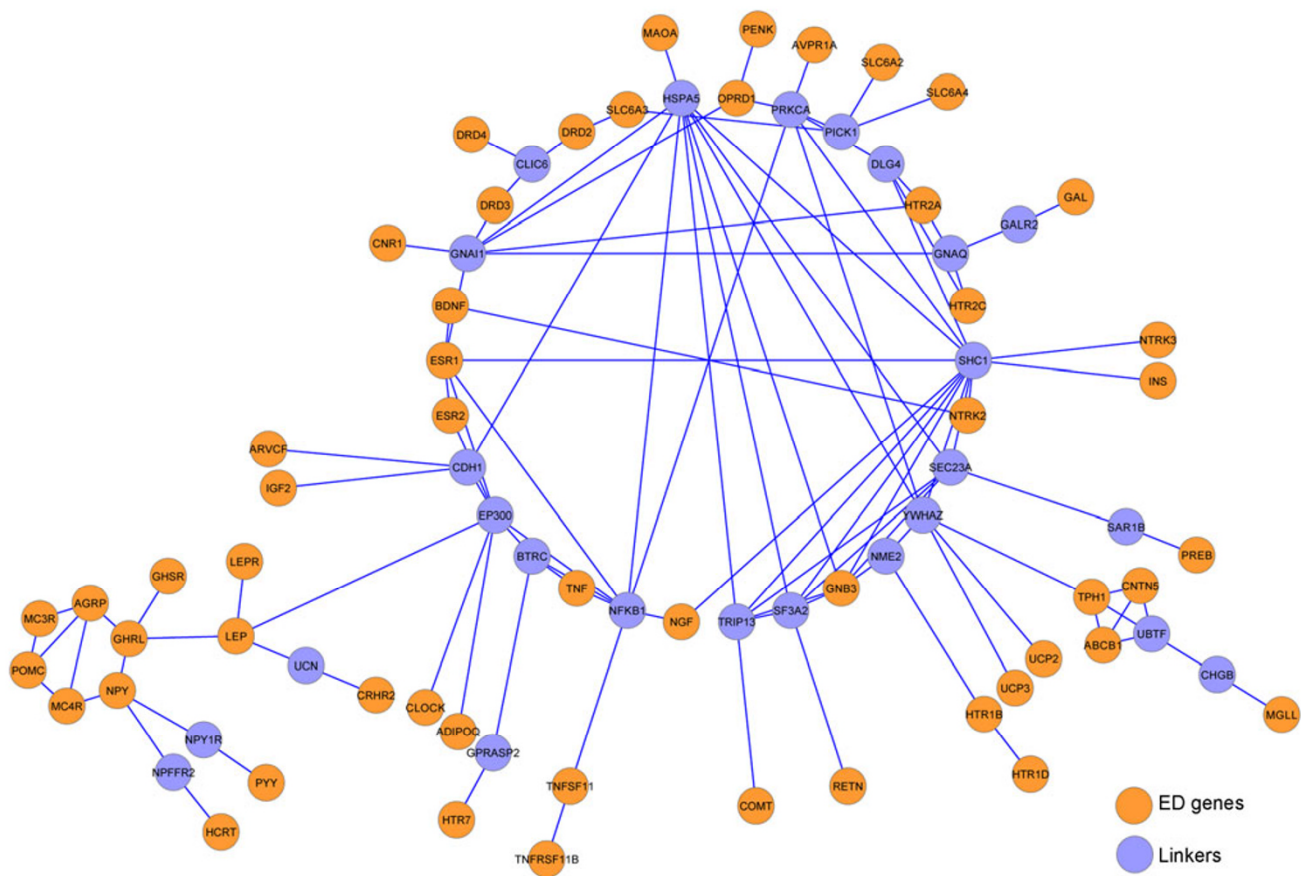


Figure 4 Reconstructed gene map for eating disorder from protein-protein interaction data. The 53 genes in orange are genes from the core dataset in our EDdb. The remaining 25 genes in blue are linker genes that bridge the 53 genes.

Table 2 Six significantly enriched signaling pathways of 78 genes in a reconstructed sub-network using PPI data^{a)}

Pathway	Source	<i>P</i> -value [*]
Neuroactive ligand-receptor interaction	KEGG	2.09×10^{-5}
Serotonergic synapse	KEGG	1.87×10^{-4}
Cocaine addiction	KEGG	6.17×10^{-4}
Dopaminergic synapse	KEGG	7.69×10^{-4}
Adipocytokine signaling pathway	KEGG	1.05×10^{-3}
Visceral fat deposits and the metabolic syndrome	BioCarta	3.52×10^{-2}

a) *, Corrected by Benjamin-Hochberg.

tral role of adipocytokine signaling in eating disorder, but also provide more clues on signaling pathways related to other neurotransmitters.

3 Discussion

The high genetic heterogeneity of eating disorder poses an enormous challenge in understanding the disease etiology. Classical identification of candidate genes for complex disorders has been conducted experimentally. However, these individual studies for identifying specific genes/variants predisposed to eating disorder from different technology

platforms are seldom incorporated with known biological information such as common pathways between eating disorder and other metabolic and psychiatric disorders. Heterogeneous data source integration might remove bias resulting from any single technology platform, and systems biology approach could provide additional insight into the genetic etiology not observed by any individual study [61–64]. Our EDdb is a high-quality resource containing genetic data from association studies, linkage scans, gene expression, literature, Gene Ontology (GO) annotations, and metabolic and signaling pathways. Each confirmed gene in our core dataset is linked with highly referred citations. By integrating protein-protein interactions and related disease data, we

expanded the candidate genes for eating disorder to broader molecular networks.

Our collection shows that 59 genes have been confirmed to be linked to eating disorder. Moreover, our systematic pathway and sub-network reconstruction revealed that the adipocytokine signaling pathway may be the central signaling pathway in eating disorder. Combining the upstream signaling events in the BioCarta pathway of “visceral fat deposits and the metabolic syndrome” with the KEGG “adipocytokine signaling pathway”, we propose a key joint signaling map for weight regulation in eating disorder, which may serve as an interesting hypothesis for further experimental testing. For instance, from the upstream signaling molecules of the “visceral fat deposits and the metabolic syndrome” pathway, only *HSD11B2* is in EXP3 with indirect evidence, while all the remaining seven genes have literature support. *HSD11B2* is the enzyme that regulates the glucocorticoid metabolism. The interaction between glucocorticoid and insulin is a very important signal for regulating long-term energy metabolism [64]. In particular, *NR3C1*, the downstream gene of *HSD11B2*, has been reported to be associated with glucocorticoid resistance, eating disorder and obesity [58,66]. Therefore, it is highly possible that *HSD11B2* is related to eating disorder. Other highly potential genes include four genes (*MTOR*, *JNK*, *IKK* and *PRKCQ*), which are absent from EDdb. These genes form a connected module centered around *PRKCQ* in response to the diacylglycerol signal and ensure cell survival and proliferation [67]. It has been reported that mouse *Prkcq* affects fat-induced insulin resistance and glucose transport [68], thus it is possible that the human *PRKCQ* is involved in insulin resistance and may also be involved in eating disorder.

Comparison of pathway contents demonstrated that eating disorder is strongly associated with obesity, schizophrenia and bipolar disorder. All four upstream signals (*TNF*, *LEP*, *LEPR* and *ADIPO*) from our core dataset and EXP1 dataset are shared with obesity, and five of the six downstream effector molecules are also shared with obesity. In our key signaling map, 25 of the 43 genes are associated with obesity. Such a high proportion of obesity-related genes in the map may provide clues for explaining some features of weight regulation in both diseases related to unhealthy lifestyle.

An interesting point to be emphasized is that the majority of the molecules in our proposed key signaling pathway are functional elements such as transcriptional regulators related to neuronal and synaptic function or key enzymes in metabolism. Therefore, the final pathway size underlying the etiology of eating disorder tends to be larger when more involved molecules are added. In summary, the level of complexity of the disorder stems from its functional components such as transcriptional regulators and key enzymes.

In conclusion, these analyses demonstrate that EDdb is a valuable tool for discovering potential candidate genes and

pathways of eating disorder, as well as potential cross-talk between obesity and eating disorder. As a result of systematically collected and extended candidate genes, EDdb may also be helpful in explaining the underlying molecular mechanisms of eating disorder.

4 Conclusion

EDdb was constructed as a free database and analysis server to enable biochemists to quickly find summarized genes related to eating disorder. Our systematic pathway studies revealed that a key joint pathway combining the KEGG “adipocytokine signaling pathway” and the BioCarta pathway of “visceral fat deposits and the metabolic syndrome” may be the central signaling pathway of weight regulation, linking eating disorder with other relevant diseases such as obesity. The EDdb database can be accessed at <http://eddb.cbi.pku.edu.cn>.

This work was supported by the National High Technology Research and Development Program of China (2008BAI64B01) and the National Natural Science Foundation of China (31171270).

- 1 Bulik C M, Hebebrand J, Keski-Rahkonen A, et al. Genetic epidemiology, endophenotypes, and eating disorder classification. *Int J Eat Disord*, 2007, 40: S52–60
- 2 Swanson S A, Crow S J, Le Grange D, et al. Prevalence and correlates of eating disorders in adolescents: Results from the national comorbidity survey replication adolescent supplement. *Arch Gen Psychiatry*, 2011, 68: 714–723
- 3 Wang K, Zhang H, Bloss C S, et al. A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol Psychiatry*, 2010, 16: 949–959
- 4 Hudson J I, Hiripi E, Pope H G Jr., et al. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol Psychiatry*, 2007, 61: 348–358
- 5 Becker K G, Barnes K C, Bright T J, et al. The genetic association database. *Nat Genet*, 2004, 36: 431–432
- 6 Chanoine P, Spector N D. Hyperlipidemia, eating disorders, and smoking cessation. *Curr Opin Pediatr*, 2008, 20: 734–739
- 7 Anzengruber D, Klump K L, Thornton L, et al. Smoking in eating disorders. *Eat Behav*, 2006, 7: 291–299
- 8 Morgan J F, Lacey J H. Smoking, eating disorders, and weight control. *Postgrad Med J*, 1999, 75: 127
- 9 Dean E, Lomi C, Bruno S, et al. Addressing the common pathway underlying hypertension and diabetes in people who are obese by maximizing health: The ultimate knowledge translation gap. *Int J Hypertens*, 2011, 2011: 835805
- 10 Schwartz M B, Henderson K E. Does obesity prevention cause eating disorders? *J Am Acad Child Adolesc Psychiatry*, 2009, 48: 784–786
- 11 Waadegaard M, Davidsen M, Kjoller M. Obesity and prevalence of risk behaviour for eating disorders among young Danish women. *Scand J Public Health*, 2009, 37: 736–743
- 12 Day J, Ternouth A, Collier D A. Eating disorders and obesity: two sides of the same coin? *Epidemiol Psychiatr Soc*, 2009, 18: 96–100
- 13 Nishisono F M. Understanding of and therapy for eating disorders-psychiatric views. *Seishin Shinkeigaku Zasshi*, 2009, 111: 854–858
- 14 Lee S, Ng K L, Kwok K, et al. The changing profile of eating disorders at a tertiary psychiatric clinic in Hong Kong (1987–2007). *Int J Eat Disord*, 2010, 43: 307–314
- 15 Javaras K N, Pope H G, Lalonde J K, et al. Co-occurrence of binge

- eating disorder with psychiatric and medical disorders. *J Clin Psychiatry*, 2008, 69: 266–273
- 16 Blinder B J. Eating disorders in psychiatric illness. *West J Med*, 1991, 155: 519–520
 - 17 Hinney A, Nguyen T T, Scherag A, et al. Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (*FTO*) variants. *PLoS ONE*, 2007, 2: e1361
 - 18 Scuteri A, Sanna S, Chen W M, et al. Genome-wide association scan shows genetic variants in the *FTO* gene are associated with obesity-related traits. *PLoS Genet*, 2007, 3: e115
 - 19 Pinheiro A P, Bulik C M, Thornton L M, et al. Association study of 182 candidate genes in anorexia nervosa. *Am J Med Genet B Neuropsychiatr Genet*, 2010, 153B: 1070–1080
 - 20 Scherag S, Hebebrand J, Hinney A. Eating disorders: The current status of molecular genetic research. *Eur Child Adolesc Psychiatry*, 2010, 19: 211–226
 - 21 Stark C, Breitkreutz B J, Chatr-Aryamontri A, et al. The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res*, 2011, 39: D698–D704
 - 22 Prasad T S, Goel R, Kandasamy K, et al. Human Protein Reference Database—2009 update. *Nucleic Acids Res*, 2008, 37: D767–772
 - 23 Willis R C, Hogue C W. Searching, viewing, and visualizing data in the Biomolecular Interaction Network Database (BIND). *Curr Protoc Bioinform*, 2006, Chapter 8, Unit 8.9
 - 24 Maglott D, Ostell J, Pruitt K D, et al. Entrez Gene: Gene-centered information at NCBI. *Nucleic Acids Res*, 2011, 39: D52–57
 - 25 Hermjakob H, Fleischmann W, Apweiler R. Swissknife—‘lazy parsing’ of SWISS-PROT entries. *Bioinformatics*, 1999, 15: 771–772
 - 26 Croft D, O’Kelly G, Wu G, et al. Reactome: A database of reactions, pathways and biological processes. *Nucleic Acids Res*, 2011, 39: D691–697
 - 27 Feramisco J D, Sadreyev R I, Murray M L, et al. Phenotypic and genotypic analyses of genetic skin disease through the Online Mendelian Inheritance in Man (OMIM) database. *J Invest Dermatol*, 2009, 129: 2628–2636
 - 28 Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics: New developments in KEGG. *Nucleic Acids Res*, 2006, 34: D354–357
 - 29 Nishimura D. BioCarta. Biotech Softw Internet Rep, 2001, 2: 117–120
 - 30 Kong L, Cheng L, Fan L Y, et al. IQdb: An intelligence quotient score-associated gene resource for human intelligence. *Database (Oxford)*, 2013, 2013: bat063
 - 31 Zhao M, Qu H. PathLocdb: A comprehensive database for the subcellular localization of metabolic pathways and its application to multiple localization analysis. *BMC Genomics*, 2010, 11: S13
 - 32 Zhao M, Chen X, Gao G, et al. RLEdb: A database of rate-limiting enzymes and their regulation in human, rat, mouse, yeast and *E. coli*. *Cell Res*, 2009, 19: 793–795
 - 33 Zhao M, Chen Y, Qu D, et al. TSdb: A database of transporter substrates linking metabolic pathways and transporter systems on a genome scale via their shared substrates. *Sci China Life Sci*, 2011, 54: 60–64
 - 34 Xie C, Mao X, Huang J, et al. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*, 2011, 39: W316–322
 - 35 Huang da W, Sherman B T, Lempicki R A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, 2009, 37: 1–13
 - 36 Bialecka M, Drozdziak M, Klodowska-Duda G, et al. The effect of monoamine oxidase B (MAOB) and catechol-O-methyltransferase (COMT) polymorphisms on levodopa therapy in patients with sporadic Parkinson’s disease. *Acta Neurol Scand*, 2004, 110: 260–266
 - 37 Urwin R E, Nunn K P. Epistatic interaction between the monoamine oxidase A and serotonin transporter genes in anorexia nervosa. *Eur J Hum Genet*, 2005, 13: 370–375
 - 38 Glaab E, Baudot A, Krasnogor N, et al. Extending pathways and processes using molecular interaction networks to analyse cancer genome data. *BMC Bioinformatics*, 2010, 11: 597
 - 39 Lorincz A M, Sukumar S. Molecular links between obesity and breast cancer. *Endocr Relat Cancer*, 2006, 13: 279–292
 - 40 Hoda M R, Hamza A, Fischer K, et al. Obesity as a risk factor for prostate cancer: Role for adipocytokines and involvement of tyrosine kinase pathway. *Aktuelle Urol*, 2010, 41: 178–183
 - 41 Nakajima T E, Yamada Y, Hamano T, et al. Adipocytokines as new promising markers of colorectal tumors: Adiponectin for colorectal adenoma, and resistin and visfatin for colorectal cancer. *Cancer Sci*, 2010, 101: 1286–1291
 - 42 Smith M R, Lee H, Fallon M A, et al. Adipocytokines, obesity, and insulin resistance during combined androgen blockade for prostate cancer. *Urology*, 2008, 71: 318–322
 - 43 Kaur T, Zhang Z F. Obesity, breast cancer and the role of adipocytokines. *Asian Pac J Cancer Prev*, 2005, 6: 547–552
 - 44 Housa D, Housova J, Vernerova Z, et al. Adipocytokines and cancer. *Physiol Res*, 2006, 55: 233–244
 - 45 Housova J, Anderlova K, Krizova J, et al. Serum adiponectin and resistin concentrations in patients with restrictive and binge/purge form of anorexia nervosa and bulimia nervosa. *J Clin Endocrinol Metab*, 2005, 90: 1366–1370
 - 46 Tagami T, Satoh N, Usui T, et al. Adiponectin in anorexia nervosa and bulimia nervosa. *J Clin Endocrinol Metab*, 2004, 89: 1833–1837
 - 47 Morton G J, Cummings D E, Baskin D G, et al. Central nervous system control of food intake and body weight. *Nature*, 2006, 443: 289–295
 - 48 Belgardt B F, Okamura T, Bruning J C. Hormone and glucose signalling in POMC and AgRP neurons. *J Physiol*, 2009, 587: 5305–5314
 - 49 de Piano A, Tock L, Carnier J, et al. The role of nutritional profile in the orexigenic neuropeptide secretion in nonalcoholic fatty liver disease obese adolescents. *Eur J Gastroenterol Hepatol*, 2010, 22: 557–563
 - 50 Speliotes E K, Willer C J, Berndt S I, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*, 2010, 42: 937–948
 - 51 Barth S, Klein P, Horbach T, et al. Expression of neuropeptide Y, omentin and visfatin in visceral and subcutaneous adipose tissues in humans: Relation to endocrine and clinical parameters. *Obes Facts*, 2010, 3: 245–251
 - 52 Maury E, Brichard S M. Adipokine dysregulation, adipose tissue inflammation and metabolic syndrome. *Mol Cell Endocrinol*, 2010, 314: 1–16
 - 53 Galic S, Oakhill J S, Steinberg G R. Adipose tissue as an endocrine organ. *Mol Cell Endocrinol*, 2010, 316: 129–139
 - 54 Benton C R, Holloway G P, Han X X, et al. Increased levels of peroxisome proliferator-activated receptor gamma, coactivator 1 alpha (PGC-1alpha) improve lipid utilisation, insulin signalling and glucose transport in skeletal muscle of lean and insulin-resistant obese Zucker rats. *Diabetologia*, 2010, 53: 2008–2019
 - 55 Phillips C M, Goumidi L, Bertrais S, et al. ACC2 gene polymorphisms, metabolic syndrome, and gene-nutrient interactions with dietary fat. *J Lipid Res*, 2010, 51: 3500–3507
 - 56 Steinberg G R, McAinch A J, Chen M B, et al. The suppressor of cytokine signaling 3 inhibits leptin activation of AMP-kinase in cultured skeletal muscle of obese humans. *J Clin Endocrinol Metab*, 2006, 91: 3592–3597
 - 57 Karp P D, Ouzounis C A, Moore-Kochlacs C, et al. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Res*, 2005, 33: 6083–6089
 - 58 Cellini E, Castellini G, Ricca V, et al. Glucocorticoid receptor gene polymorphisms in Italian patients with eating disorders and obesity. *Psychiatr Genet*, 2010, 20: 282–288
 - 59 Dostalova I, Kunesova M, Duskova J, et al. Adipose tissue resistin levels in patients with anorexia nervosa. *Nutrition*, 2006, 22: 977–983
 - 60 Krizova J, Dolinkova M, Lacinova Z, et al. Adiponectin and resistin gene polymorphisms in patients with anorexia nervosa and obesity and its influence on metabolic phenotype. *Physiol Res*, 2008, 57: 539–546

- 61 Zhao M, Qu H. High similarity of phylogenetic profiles of rate-limiting enzymes with inhibitory relation in human, mouse, rat, budding yeast and *E. coli*. *BMC Genomics*, 2011, 12: S10
- 62 Zhao M, Qu H. Human liver rate-limiting enzymes influence metabolic flux via branch points and inhibitors. *BMC Genomics*, 2009, 10: S31
- 63 Edwards Y J, Beecham G W, Scott W K, et al. Identifying consensus disease pathways in Parkinson's disease using an integrative systems biology approach. *PLoS ONE*, 2011, 6: e16917
- 64 Zhao M, Sun J, Zhao Z. Synergetic regulatory networks mediated by oncogene-driven microRNAs and transcription factors in serous ovarian cancer. *Mol BioSyst*, 2013, 9: 3187–3198
- 65 Andrews R C, Walker B R. Glucocorticoids and insulin resistance: Old hormones, new targets. *Clin Sci (Lond)*, 1999, 96: 513–523
- 66 Rosmond R, Chagnon Y C, Holm G, et al. A glucocorticoid receptor gene marker is associated with abdominal obesity, leptin, and dysregulation of the hypothalamic-pituitary-adrenal axis. *Obes Res*, 2000, 8: 211–218
- 67 Meller N, Altman A, Isakov N. New perspectives on PKC θ , a member of the novel subfamily of protein kinase C. *Stem Cells*, 1998, 16: 178–192
- 68 Kim J K, Fillmore J J, Sunshine M J, et al. PKC- θ knockout mice are protected from fat-induced insulin resistance. *J Clin Invest*, 2004, 114: 823–827

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Supporting Information

Table S1 Top 10 statistically significant enriched signaling pathways from the KEGG and BioCarta databases for different expanded datasets in EDdb

The supporting information is available online at life.scichina.com and www.springerlink.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.